REVIEW

Mechanism of action, potency and efficacy: considerations for cell therapies

Carl G. Simon Jr.[1*] , Erich H. Bozenhardt[2], Christina M. Celluzzi[3], David Dobnik[4,5], Melanie L. Grant[6], Uma Lakshmipathy[7], Thiana Nebel[8], Linda Peltier[9], Anthony Ratcliffe[10], James L. Sherley[11], Glyn N. Stacey[12,13,14], Rouzbeh R. Taghizadeh[15], Eddie H. P. Tan[16] and Sandrine Vessillier[17]

Abstract

One of the most challenging aspects of developing advanced cell therapy products (CTPs) is defining the mechanism of action (MOA), potency and efficacy of the product. This perspective examines these concepts and presents helpful ways to think about them through the lens of metrology. A logical framework for thinking about MOA, potency and efficacy is presented that is consistent with the existing regulatory guidelines, but also accommodates what has been learned from the 27 US FDA-approved CTPs. Available information regarding MOA, potency and efficacy for the 27 FDA-approved CTPs is reviewed to provide background and perspective. Potency process and efficacy process charts are introduced to clarify and illustrate the relationships between six key concepts: MOA, potency, potency test, efficacy, efficacy endpoint and efficacy endpoint test. Careful consideration of the meaning of these terms makes it easier to discuss the challenges of correlating potency test results with clinical outcomes and to understand how the relationships between the concepts can be misunderstood during development and clinical trials. Examples of how a product can be "potent but not efficacious" or "not potent but efficacious" are presented. Two example applications of the framework compare how MOA is assessed in cell cultures, animal models and human clinical trials and reveals the challenge of establishing MOA in humans. Lastly, important considerations for the development of potency tests for a CTP are discussed. These perspectives can help product developers set appropriate expectations for understanding a product's MOA and potency, avoid unrealistic assumptions and improve communication among team members during the development of CTPs.

Keywords Cell therapy product, Efficacy, Efficacy endpoint, Efficacy endpoint test, Mechanism of action, Potency, Potency test

Introduction

Challenges of MOA and potency

Determining the MOA and developing adequate potency tests remain challenging for CTPs [1–6]. US regulations require that CTPs have a potency test for licensure. Potency tests should be based on the product's MOA [7, 8]. The goal of a potency test is to assure that the product is able to achieve its intended mechanism of action.

Other major roles of potency tests are to assess manufacturing consistency and product stability. Potency tests are often bioassays that involve measuring a response in cells.

For many of the 27 US FDA-approved CTPs (as of February 2024), the relationships between the potency tests and proposed MOAs are unclear. Table 1 highlights representative examples of the information regarding potency and MOA taken from the regulatory documentation for seven of the FDA-approved CTPs (see Supplementary File 1 for information about all 27 CTPs). The FDA has a useful website containing

*Correspondence:
Carl G. Simon Jr.
carl.simon@nist.gov
Full list of author information is available at the end of the article

Table 1 Regulatory comments regarding potency & MOA For FDA-approved cell therapies

Product (year approved, company): Description	Indication	Potency test	MOA
Provenge (2011, Dendreon): Autologous cellular immunotherapy (CD54+ cells activated with PAP-GM-CSF)	Prostate cancer	Surface expression of CD54 (ICAM-1) on APCs after culture with PAP-GM-CSF (flow cytometry)	*Package Insert:* "While the precise **mechanism of action is unknown**, PROVENGE is designed to induce an immune response targeted against PAP, an antigen expressed in most prostate cancers."
Gintuit (2012, Organogenesis): Allogeneic cultured keratinocytes and fibroblasts in bovine collagen	Oral mucogingival defects	Histology with morphological assessments	*Package Insert:* "GINTUIT does not function as a tissue graft. The MOA by which GINTUIT increases keratinized tissue at the treated site has not been identified." *FDA SBRA:* "The histology assay is a good measure of the structural integrity of the product, however, **the assay is not an adequate, sensitive measure of biological activity.**"
MACI (2016, Vericel): Autologous cultured chondrocytes on porcine collagen membrane	Knee cartilage defects	PCR measurement of aggrecan gene expression	*Package Insert:* "No clinical pharmacology studies have been conducted with MACI and a **mechanism of action has not been established**."
Kymriah (2017, Novartis): CD19-directed genetically modified autologous T-cell immunotherapy	B-cell acute lymphoblastic leukemia	Interferon-γ production by test article upon stimulation with CD19+ cells	*Package Insert:* "Upon binding to CD19-expressing cells, the CAR transmits a signal to promote T-cell expansion, activation, target cell elimination, and persistence of the KYMRIAH cells." *FDA Briefing Document:* "In the clinical trials, IFN-γ production varied greatly from lot-to-lot, **making it difficult** to correlate IFN-γ production with in vitro safety or efficacy."
Stratagraft (2021, Stratatech): Allogeneic cultured keratinocytes and dermal fibroblasts in murine collagen	Thermal burns	*Not reported*	*Package Insert:* "In vitro studies have shown that Stratagraft **secretes** human growth factors and cytokines and contains human ECM proteins."
Rethymic (2021, Enzyvant): Allogeneic processed thymus tissue	Athymia	Histology	*Package Insert:* "The **proposed** mechanism of action involves the migration of recipient T cell progenitors from the bone marrow to the implanted Rethymic slices, where they develop into naive immunocompetent recipient T cells."
Lantidra (2023, CellTrans): Allogeneic pancreatic islets	Type 1 diabetes	i) insulin release in glucose stimulated islets; ii) islet yield by microscopy; iii) viability by microscopy	*Package Insert:* "The primary mechanism of action of LANTIDRA is **believed** to be secretion of insulin by infused (transplanted) β- cells."
Amtagvi (2024, Iovance): Autologous T-cell therapy made of ex vivo-expanded lymphocytes harvested from patient tumors	Melanoma	i) redacted, ii) redacted, iii) redacted, iv) redacted, v) dose (total viable cells), vi) redacted and vii) redacted	*Package Insert:* "The specific mechanism of action of AMTAGVI (lifileucel) is **unknown**."

FDA Food and Drug Administration, *PAP-GM-CSF* human prostatic acid phosphatase (PAP), an antigen expressed in prostate cancer tissue, linked to human granulocyte–macrophage colony-stimulating factor (GM-CSF), an immune cell activator, *SBRA* Summary Basis for Regulatory Action. Bolded text highlights areas of uncertainty or concern and bolding was added by the authors

Source: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/approved-cellular-and-gene-therapy-products

Simon Jr. *et al. Journal of Translational Medicine*   (2024) 22:416

Page 3 of 17

information for approved CTPs [9]. Two documents posted for each product are particularly useful: the "Product Insert" and "Summary Basis for Regulatory Action" (SBRA), which contain information about the clinical trials, product testing, potency and MOA. The regulatory documentation for Provenge, Gintuit, MACI and Amtagvi indicate that the MOA is not known for these products (Table 1). For Kymriah, the documentation states that it is difficult to correlate the potency test results with efficacy. The documentation for Stratagraft discusses data regarding the product activity but does not directly state whether this activity is to be regarded as the MOA. For Rethymic and Lantidra, the documentation uses the words "proposed" and "believed" when discussing the MOA to indicate uncertainty. The lack of clarity regarding potency and MOA lead to the question: "Why is there such a challenge with MOA and potency?" This perspective article attempts to shed light on this question.

Table 2 summarizes information on potency tests for the 27 approved CTPs. Supplementary File 1 summarizes the key aspects of each of the FDA-approved CTPs including product name, year approved, sponsor, product description, indication, clinical trial structure, efficacy endpoints, MOA, potency test, comments and references. This information was used as background for many of the points highlighted in this perspective.

Correlation of potency test with clinical outcome

It is desirable for the potency test to reflect clinical efficacy [7, 8]. Manufactured units that show high potency via the potency test should tend to show an efficacious benefit to patients, while units with lower potency, as measured by the potency test, should have a less efficacious benefit to patients. However, a correlation between the potency test and clinical efficacy is not required. If a product is efficacious and the risk–benefit profile is acceptable, then a product may receive regulatory marketing clearance even if the potency test does not correlate with efficacy endpoint test results (Table 1). For CTPs, MOAs may not be fully understood making it difficult to relate MOAs to potency or efficacy.

Kymriah: relationship between potency test results and clinical outcome

In 2017, Kymriah (tisagenlecleucel) was the first chimeric antigen receptor (CAR) T-cell therapy that was cleared for marketing by FDA. Kymriah is used to treat leukemia and potency was defined as the ability of the CAR T-cells to secrete interferon-γ (IFN-γ) following exposure to target cells expressing CD19. The FDA held a meeting of the Oncologic Drugs Advisory Committee to discuss this revolutionary technology in a public forum. The

online documentation contained a noteworthy graph showing the relationship between potency measurements and efficacy endpoints (Fig. 1) [10]. To the authors' knowledge, this is the only publicly available data to show the relationship between potency testing results and clinical outcome for a US-approved CTP. The graph shows that the potency test results correlated with remission, but that there was overlap between responders and non-responders. The FDA Briefing Document from the meeting stated "In the clinical trials, IFN-γ production varied greatly from lot-to-lot, making it difficult to correlate IFN-γ production in vitro to tisagenlecleucel safety or efficacy" (Table 1) [11].

Definitions

Clear definitions are critical to improving the understanding of the relationships among MOA, potency and efficacy. Figure 2a provides definitions of 6 key terms and delineates their relationships as discussed below. Definitions for terms were adapted from existing definitions but may not be identical to definitions in the source documents. Definitions were created for terms that had not previously been defined.

Measurand and measurement

Since the current discussion is focused on a product attribute (potency) and how to measure it (potency test), it is important to understand the key metrological term, measurand.

Measurand: "*the quantity or property intended to be measured*" [12]

A measurand is the attribute or property of a material that is being assessed. A measurand and a material attribute can be the same thing when the goal of a measurement is to measure the material attribute. The use of the word "intended" in this definition is intentional and important. This is because it may be impossible to measure what one intends to measure. Further, due to experimental artifacts, what was measured during the measurement may not be what was intended [13]. These points are reminders that there are no perfect measurements and that all measurements have false positives and false negatives. This concept is key to misunderstandings surrounding potency tests.

It is important to independently define the "measurand" (the attribute being measured) and the "measurement", according to the International Vocabulary of Metrology (VIM) [12].

Measurement: "process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity" [12]

Table 2 Potency tests and clinical trial designs for the 27 US-approved therapies containing live human cells

Type	Product (year)	Potency Tests (from FDA Summary Basis for Regulatory Action)	Clinical Trial Design
HSCs (allogeneic hematopoietic progenitor cell therapies)	Hemacord (2011)	*FDA SBRA:* i) Total nucleated cell number; ii) viability of CD45+ cells; iii) viable CD34+ cell count; iv) CFU	Single-arm
	Clinimmune (2012)	*FDA SBRA:* i) Total nucleated cells; ii) viability of total nucleated cells; iii) viable CD34+ cell count; iv) redacted	Single-arm
	Ducord (2012)	*FDA SBRA:* i) Total nucleated cells; ii) viable nucleated cells; iii) viable CD34+ cells (flow cytometry); iv) redacted; v) redacted	Single-arm
	Lifesouth (2013)	*FDA SBRA:* i) Total nucleated cells; ii) viable nucleated cells; iii) viable CD34+ cells (flow cytometry); iv) redacted	Single-arm
	Bloodworks (2016)	*FDA SBRA:* i) Total nucleated cells; ii) viable nucleated cells; iii) viable CD34+ cells (flow cytometry); iv) redacted	Single-arm
	Allocord (2016)	*FDA SBRA:* i) Total nucleated cells; ii) viable nucleated cells; iii) viable CD34+ cell count; iv) CFU	Single-arm
	Clevecord (2016)	*FDA SBRA:* i) Total nucleated cell number; ii) viable nucleated cells; iii) viable CD34+ cell count; iv) redacted	Single-arm
	MD Anderson (2018)	*FDA SBRA:* i) Total CD34+ count; ii) total nucleated cell count; iii) nucleated red blood cell; iv) viability of nucleated cells; v) viable CD34+ cells; vi) CFU assay	Single-arm
CAR T-cells (chimeric antigen receptor T-cell therapy)	Kymriah (2017)	*FDA Panel Meeting Slides:* i) Determination of CAR expression by flow cytometry; ii) Interferon-γ production by product upon stimulation with CD19+ cells	Single-arm
	Yescarta (2017)	*FDA SBRA:* i) Cell viability; ii) anti-CD19 CAR expression; iii) redacted *Papadouli et al., 2020:* Interferon-γ production by product upon stimulation with CD19+ cells	Single-arm
	Tecartus (2020)	*FDA SBRA:* i) Cell viability; ii) anti-CD19 CAR expression; iii) redacted	Single-arm
	Breyanzi (2021)	*FDA SBRA:* i) Redacted	Single-arm with 3 dose cohorts
	Abecma (2021)	*FDA SBRA:* i) Redacted *EMA Assessment Report:* Interferon-γ production by product upon stimulation with BCMA+ cells	Single-arm with 3 dose cohorts
	Carvykti (2023)	*FDA SBRA*: i) CAR expression from viable T cells; ii) redacted	Single-arm
Other	Provenge (2010)	*FDA SBRA:* i) Number of CD54+ cells (flow cytometry); ii) increased expression of CD54 on the surface of antigen presenting cells after culture with PAP-GM-CSF (flow cytometry)	Two-arm (compared to unactivated autologous PBMCs)
	Laviv (2011)	*FDA SBRA:* i) Cell count; ii) cell viability; iii) collagen production by the cells	Two-arm (compared to placebo)
	Gintuit (2012)	*FDA SBRA:* Histology with morphological assessments: epidermal coverage, epidermal development, basal cell layer keratinocyte viability, suprabasal cell layer keratinocyte viability, dermal thickness, fibroblast density, and matrix integrity	Two-arm (compared to standard of care)
	Maci (2016)	*FDA SBRA:* i) Cell number; ii) redacted; iii) redacted *Rapko et al., 2007:* PCR measurement of aggrecan gene expression	Two-arm (compared to standard of care)
	Stratagraft (2021)	*FDA SBRA:* Redacted	Two-arm (compared to standard of care)
	Rethymic (2021)	*FDA SBRA:* Histology-based (for tissue organization, viability and retention of important cell types believed to be important for product function)	Single-arm
	Zynteglo (2022)	*FDA SBRA:* i) Vector copy number (qPCR); ii) percent LVV+ cells; iii) colony forming cells; iv) βA-T87Q-globin quantitative protein expression; v) redacted; vi) redacted	Single arm (compared to symptoms prior to treatment)
	Skysona (2022)	*FDA SBRA:* i) Vector copy number (qPCR); ii) percent LVV+ cells; iii) percent ADLP+ cells; iv) redacted; v) redacted; vi) redacted	Single arm
	Omisirge (2023)	*FDA SBRA:* CD34+ cell fold-increase	Two-arm (compared to umbilical cord blood transplantation)
	Lantidra (2023)	*FDA Advisory Committee Meeting:* i) ELISA quantification of insulin release in glucose stimulated islets; ii) islet yield assessed by dithizone stain and microscopic quantification; iii) viability assessed by SYTO 13 green/ethidium bromide staining and microscopic evaluation	Single-arm
	Casgevy (2023)	*FDA SBRA:* i) On-target editing frequency (tracking of indels by decomposition, TIDE); ii) redacted; iii) redacted.	Single-arm
	Lyfgenia (2023)	*FDA SBRA:* i) Vector copy number; ii) redacted, iii) redacted; iv) redacted; v) redacted; vi) βA-T87Q-globin quantitative protein expression.	Single-arm
	Amtagvi (2024)	*FDA SBRA:* i) redacted, ii) redacted, iii) redacted, iv) redacted, v) dose (total viable cells), vi) redacted and vii) redacted.	Single-arm

Abbreviations: CFU = colony forming unit; ELISA = enzyme linked immunosorbent assay; FDA SBRA = Food and Drug Administration Summary Basis for Regulatory Action; PAP-GM-CSF = human prostatic acid phosphatase (PAP), an antigen expressed in prostate cancer tissue, linked to human granulocyte-macrophage colony-stimulating factor (GM-CSF), an immune cell activator; PBMC = peripheral blood mononuclear cells; SMN = survival motor neuron; PCR = polymerase chain reaction

There may be a variety of different methods for making a measurement of a material attribute and the results of the measurements may not agree. For example, two labs may try to make the same measurement but get different results. It may be "that different experiments are inadvertently being performed (*i.e.*, there are critical experimental differences that are not accounted for)" [14]. This is not a new challenge for cellular therapies but has existed more broadly for biomedical research.
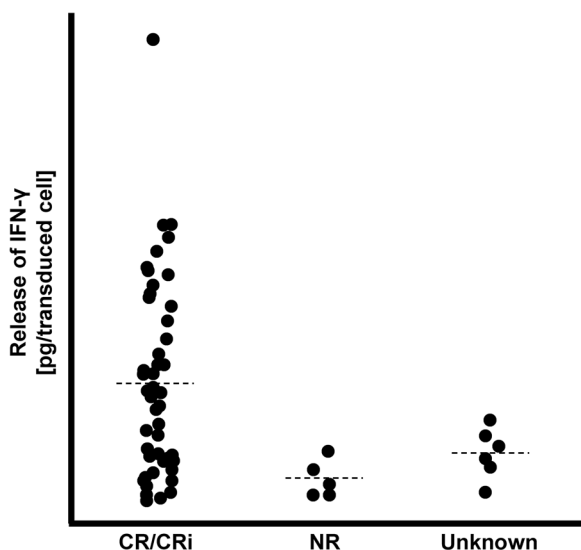
Fig. 1 Relationship between potency test results and efficacy endpoint for Kymriah clinical trial for the treatment of pediatric acute lymphoblastic leukemia after three months. The potency test was interferon-γ (IFN) production by test article upon stimulation with CD19 + cells. Data are adapted from FDA Advisory Committee Meeting materials [10]. Data are from 63 patients: 52 CR/CRi, 5 NR and 6 Unknown. CR = complete remission; CRi = complete remission with incomplete blood count recovery; NR = nonresponder; Unknown = unknown response

Mechanism of action (MOA), potency, potency test, efficacy, efficacy endpoint and efficacy endpoint test

Following are six key definitions that are charted in Fig. 2a. These definitions have been adapted from the cited sources. The source definitions are provided in Supplementary File 2 as a reference.

- *Mechanism of action (MOA):* The specific process, often pharmacologic, through which a product produces its intended effect [15, 16]
- *Potency:* The attribute of a product that enables it to achieve its intended mechanism of action [17–19]
- *Potency Test:* A test that measures the attribute of a product that enables it to achieve its intended mechanism of action [20]
- *Efficacy:* The ability of the product to have the desired effect in patients [21]
- *Efficacy Endpoint:* Attributes related to how a patient feels, functions or survives [22, 23]
- *Efficacy Endpoint Test:* A test that measures attributes related to how a patient feels, functions or survives (the authors are not aware of any existing definition for this term)

It is key that these six concepts be defined separately from one another in order to understand how they relate to one another.

Potency ≠ efficacy

A common mistake is to assume that potency and efficacy are the same thing and that a potency test that measures potency is also a measure of efficacy. This cannot be true since efficacy can only be measured by clinical response. Potency tests are laboratory assays which may or may not be a predictor of clinical response. Another way to think about it is "potency is laboratory" whereas "efficacy is clinical"; and the two are tied together by the MOA.

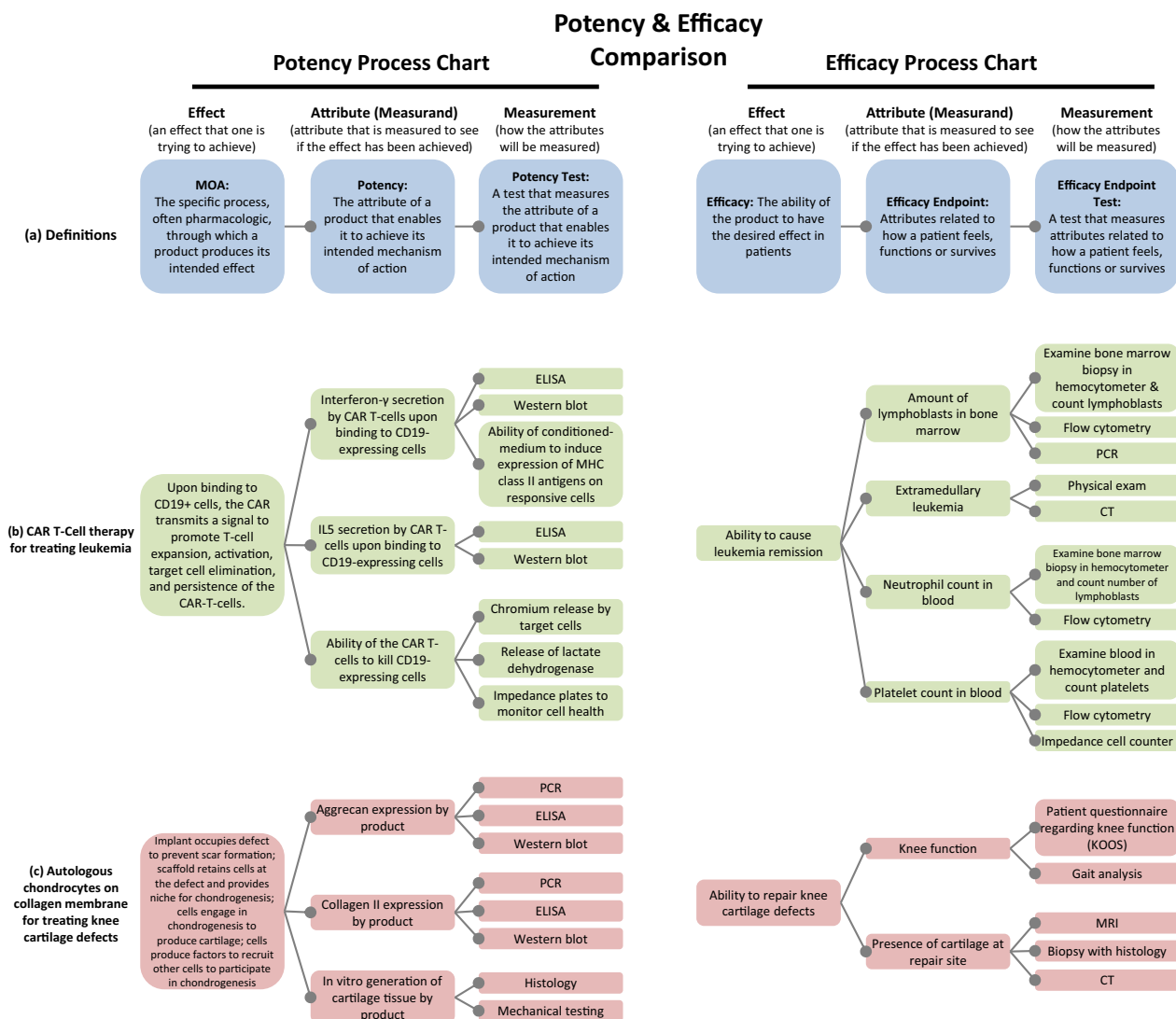### Potency process charts and efficacy process charts

The definitions given in Fig. 2a are applied to practical examples of two approved CTPs: a CAR T-cell therapy in Fig. 2b [10, 11, 24–32] (based on Kymriah) and tissue-engineered chondrocytes on a collagen membrane in Fig. 2c [33–37] (based on MACI). Each chart is composed of 2 triads: an effect, an attribute and a measurement which are applied to MOA in the first three boxes and to efficacy in the fourth, fifth and sixth boxes. Another way to think of these triads is 1) an effect that one is trying to achieve, 2) an attribute that can be measured to see if the effect has been achieved (the measurand), and 3) how the attributes will be measured (the measurement).

The goal of the potency and efficacy process charts in Fig. 2a is to distinguish and independently define each of the 6 components, making it easier to understand how their interrelationships can break down during product development and clinical trials.

Separate MOA from potency

A benefit of separating the "MOA" from the "potency" attribute is that it allows for the defined potency attribute to be incorrect. For example, potency of a CAR T-cell therapy could be defined as its ability to secrete IFN-γ upon recognizing target cells (Fig. 2b). However, it could be that IFN-γ secretion is not involved in the proposed MOA of the product (target cell elimination) and the product may have an unknown MOA.

Separation of "MOA" from the "potency" attribute also allows different biological activities related to the MOA to be defined as potency. For example, the potency of a CAR T-cell therapy could be defined as its ability to secrete IFN-γ upon binding to target cells (Fig. 2b), its ability to secrete IL5 upon binding to target cells or the ability of the CAR T-cells to kill target cells [31]. All, some or none of these could be solely indicative of the MOA. It may be that IFN-γ secretion is an essential factor in the MOA, and that IL5 and cell killing are not

Fig. 2 Potency and efficacy process charts. Comparison of potency and efficacy definitions and processes for two example products demonstrates the parallelism between the two processes. **a** Definitions of the components of a potency and efficacy process chart. **b** Potency and efficacy process charts for treating leukemia with a CAR T-cell therapy (based on Kymriah). **c** Potency and efficacy process charts for a product for treating knee cartilage defects with chondrocytes on a collagen membrane (based on MACI). The figure graphically represents the relationships among the three inherent components of any analytical assessment: effect; attribute (measurands); and measurement; for both potency and efficacy. The examples are meant to illustrate concepts and are not intended as recommendations for potency tests or efficacy endpoint tests. The charts are not comprehensive and there may be other MOAs, attributes or tests that are not mentioned. *CT* computed tomography scan, *ELISA* enzyme linked immunosorbent assay, *IL5* interleukin 5, *KOOS* knee injury and osteoarthritis outcome score, *MHC* major histocompatibility, *MRI* magnetic resonance imaging, *PCR* polymerase chain reaction. The cited sources were used to assemble the content for (**b**) Kymriah [10, 11, 24–32] and (**c**) MACI [33–37]

major contributors. Or, maybe none of these activities are part of the MOA.

Separate potency from potency test
Likewise, "potency" and "potency test" should be distinct in the event that 1) the "potency test" is not a reliable measurement of the potency attribute and 2) there are multiple ways to measure the potency attribute.

For case 1), measurement of IFN-γ secretion by CAR T-cells by enzyme linked immunosorbent assay (ELISA) (Fig. 2b) could lead to artifactual data that do not accurately measure the amount of IFN-γ released. For case 2, IFN-γ release could be measured by ELISA or Western blot and the results may not agree. These possibilities are obscured if the potency attribute is defined in terms of a specific measurement, emphasizing the value in keeping

potency (attribute or measurand) and the potency test (measurement) distinct.

Separate efficacy from efficacy endpoint

The concept of "efficacy" should remain generalized and distinct from the "efficacy endpoint", while the "efficacy endpoint" should be specific to a patient symptom or attribute. For instance, when developing a CAR T-cell therapy, developers could define "efficacy" as "ability to cause leukemia remission", which is a general definition that does not mention patient attributes or symptoms (i.e., the "efficacy endpoints") (Fig. 2b). If efficacy is defined as "amount of lymphoblasts" (instead of as the "ability to cause leukemia remission"), then developers are led to believe that "amount of lymphoblasts" is certain to be indicative of the disease status of a patient, and the possibility that "amount of lymphoblasts" is not linked to the disease state is more easily overlooked. The measurement of the "amount of lymphoblasts" in bone marrow may not be indicative of a particular blood cancer, and it may be that other efficacy endpoints should be considered, such as presence of extramedullary tumors. Having separate definitions for "efficacy" (the effect you are trying to achieve) and "efficacy endpoints" (patient attributes or symptoms) allows for the possibility that a given "efficacy endpoint" is not indicative of the disease state. This distinction is subtle but it avoids the false assumptions that lead to confusion when unexpected clinical results are observed.

Separate "efficacy endpoint" from "efficacy endpoint test"

Separation of the "efficacy endpoint" from the "efficacy endpoint test" (the measurement) makes it easier to deal with false measurement results. The VIM provides separate definitions for measurand and measurement [12], so it makes sense to do the same here. For example, the amount of lymphoblasts in bone marrow may be used as an efficacy endpoint for leukemia treatment (Fig. 2b). The measurement of the amount of lymphoblasts could be conducted by hemocytometer counting (efficacy endpoint test) but the results could be deceptive. There may be an interference in the hemocytometer count (e.g., cell clumping, debris) causing it to give false readings. The efficacy endpoint should be defined without referring to a specific measurement method, so that the possibility of false results is not obscured.

Separation of the "efficacy endpoint" from "efficacy endpoint test" makes it easier to deal with different measurements of an attribute. The number of lymphoblasts in marrow could be measured by counting in a hemocytometer, by PCR (polymerase chain reaction) or by flow cytometry; and the three measurements may not agree [30]. If the efficacy endpoint is defined as a hemocytometer measurement of lymphoblasts, then the circumstance of multiple measurements that may not agree becomes intractable.

"Potent but not efficacious" and "not potent but efficacious"

Since potency and efficacy cannot be the same thing (because potency is measured by a lab test and efficacy is measured by clinical trial), then it follows that a therapeutic could be "potent but not efficacious" or "not potent but efficacious." These are confusing scenarios, but it is important to consider them, since they get to the heart of the challenge surrounding MOA, potency and efficacy.

Potent but not efficacious: wrong patient population

An example of "potent but not efficacious" is giving a potent chemotherapeutic to a patient with bronchitis. Imagine that the product is truly potent for its intended MOA and is truly efficacious when used as intended in the treatment of cancer. Yet, the chemotherapeutic will not help the patient with bronchitis. This could be a case of a misdiagnosis or treating the wrong patient population.

This example could seem irrelevant, but identifying the appropriate indication and selecting the appropriate patient inclusion and exclusion criteria for a clinical trial are critical. When statistical significance is not observed in a costly and lengthy clinical trial, there are often controversial post-hoc subgroup analyses that identify responsive cohorts. A better understanding of MOA helps with identification of patients most likely to respond prior to starting a clinical trial. An example is companion diagnostics [38], such as the well-known example of breast cancer screening to treating HER2/*neu*-positive patients with Herceptin, an anti-HER2/*neu* antibody drug [39]. The FDA website currently lists 169 companion diagnostics, which highlights the value of identifying the correct patient population [38].

Another example is the indication for Kymriah, which is carefully worded with 5 qualifiers to specifically identify those patients most likely to benefit from treatment: "patients up to 25 years of age with B-cell precursor acute lymphoblastic leukemia (ALL) that is refractory or in second or later relapse" [28] (Additional file 1). Another example is the carefully worded indication for Lantidra which has six qualifiers: "The treatment of adults with Type 1 diabetes who are unable to approach target HbA1c because of current repeated episodes of severe hypoglycemia despite intensive diabetes management and education." Careful thinking and strategy are required to write

CTP indications so that patients for whom the product is most likely to be efficacious are identified.

Potent but not efficacious: incorrect hypothesis regarding the disease mechanism

Another example of "potent but not efficacious" is having an incorrect hypothesis regarding the MOA for the targeted indication. Imagine the therapeutic is truly potent and achieves its intended MOA. Yet, it does not benefit the patient population because the hypothesis regarding the MOA is wrong for the intended indication. Achieving the intended MOA does not have an effect in treating the intended indication. This exercise highlights the importance of separating the concepts of potency and efficacy. If potency and efficacy are equated to one another, then "potent but not efficacious" becomes inconceivable.

Another useful example for understanding "potent but not efficacious" is as follows. If a person takes an aspirin for a headache and the headache is not alleviated, does that mean the aspirin is not potent? The aspirin is probably still potent, inhibits cyclooxygenase and blocks prostaglandin synthesis to achieve its intended MOA [40]. Yet, the aspirin is not efficacious in curing the headache. The headache may be caused by factors unrelated to cyclooxygenase or prostaglandins.

Not potent but efficacious: alternate MOA

An example of "not potent but efficacious" could be a therapeutic that is effective due to an MOA that is not the proposed MOA. Imagine the therapeutic is truly *not potent* for its intended MOA, but the therapeutic is truly effective in treating the intended indication. It could be that the proposed MOA was incorrect, and the product is effective due to an alternate MOA (perhaps an unknown biological activity). To revisit the CAR T-cell example discussed in Fig. 2b, potency of a CAR T-cell therapy may not be due to its ability to secrete IFN-γ upon binding to target cells, but instead could be due to secretion of IL5 upon binding to target cells or the ability of the CAR T-cells to kill target cells via perforin-granzyme or Fas–Fas ligand interactions [41].

Not potent but efficacious: false negative potency test

Another example of "not potent but efficacious" is a false negative potency test result. Imagine that a therapeutic is truly potent. It achieves its intended MOA and achievement of the intended MOA is truly effective in treating the intended indication. Yet, the potency test indicates that the therapeutic is not potent. In this scenario, the potency test may be a poor measure of true potency and may yield incorrect results: that the therapeutic is not potent. In this case, the potency test is giving false negative results. For example, the potency test may not be stable and reagent degradation results in intermittent false negatives.

MOA case studies: aspirin and acetaminophen
Aspirin MOA

The MOA of aspirin (acetylsalicylic acid) is useful to consider. Aspirin is considered the most widely used drug of all time [42]. For thousands of years, plant extracts, such as from willow bark which contains salicylate, have been used to treat rheumatism (joint inflammation) and pain. In 1897, Bayer reported a synthetic version of acetylsalicylic acid that was named aspirin. In the early 1970s, research by John Vane shed light on aspirin's MOA with the discovery of cyclooxygenase (COX-1) and its acetylation by aspirin [43–45], leading to the 1982 Nobel Prize in Medicine [46]. However, other COX isoforms were discovered; COX-2 in 1991 [47] and COX-3 in 2002 [48]; and COX-3 was inhibited by aspirin. In the 2000s, additional MOAs for aspirin were elucidated: i) uncoupling of oxidative phosphorylation [49], ii) promotion of nitric oxide synthesis to inhibit inflammation [50] and iii) inhibition of NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells) activation to inhibit inflammation [51]. This evidence suggests that aspirin has multiple (and possibly multifactorial) MOAs and can affect at least 5 pathways: COX-1, COX-3, oxidative phosphorylation, nitric oxide and NF-κB. Future studies may identify new MOAs for aspirin. Thus, even after thousands of years of use, and being the most widely used drug of all time, the MOA of aspirin is still unclear.

Acetaminophen MOA

Another example of a widely used drug whose MOA is uncertain is acetaminophen (paracetamol). Acetaminophen, a small molecule drug with a molecular mass of 151 g/mol, is probably the most widely prescribed drug for children and is used for pain management and fever reduction. It was first synthesized in 1878, first used clinically in 1887, and was first marketed in the US in the 1950s [52]. Acetaminophen's package insert states that "The precise mechanism of the analgesic and antipyretic properties of acetaminophen is not established but is thought to primarily involve central actions" [53]. There is evidence that acetaminophen inhibits prostaglandin synthesis [54], inhibits COX-3 [48], activates a vanilloid receptor (TRPV1) [55] and modulates the cannabinoid system [56].

CTPs vs. "aspirin and acetaminophen"

Aspirin and acetaminophen, small molecule drugs with molecular masses of 180 g/mol and 151 g/mol, respectively, each has multiple activities. In small molecule drug discovery, the term polypharmacology means "the

binding of a drug to multiple target proteins, with clinical effects being mediated through the modulation of the set of protein targets" [57]. The precise MOA for many drugs is unknown, which indicates that knowing the MOA is desirable but unnecessary for having an effective therapeutic [58]. A cell therapy delivers millions of metabolically active cells to a patient, where each cell is a bag of thousands of different molecules, including proteins, lipids, carbohydrates and nucleic acids, and each of these molecules could have multiple activities. Given the fact that the MOAs for aspirin and acetaminophen are still being investigated, determining the precise MOA for a cell-containing therapeutic seems especially challenging. Further, without a good understanding of the MOA, it is hard to imagine that there will be good potency tests, given that potency tests should be based on the MOA. Given the tools currently at our disposal, it may be useful to consider the establishment of a product's MOA as an aspirational goal, instead of as something that can be known with certainty. This may change with development of improved technology for determining MOA.

Clinical trials are not designed to establish MOA
Example application: establishing MOA for a CAR T-cell product
It is useful to consider an example application for establishing the MOA of a CAR T-cell therapy directed against leukemia cells expressing a target surface antigen (based on Kymriah). The proposed MOA for this product is that upon binding to CD19+cells (cancer cells), the CAR transmits a signal to promote T-cell expansion, activation, target cell elimination, and persistence of the CAR-T-cells.

Table 3 shows experiments that can be used to establish the MOA. The tests can be run in cell culture using an in vitro cell killing assay, in animals using an animal model for leukemia and in a human clinical trial to treat leukemia patients. The test article is "CAR T-cells directed against leukemia cells expressing a target surface antigen" as shown in the first row of the table. The second row is a negative control, to assess if the test article has an effect above background signal. The negative control could be an untreated well for cell culture, could be a sham CAR T-cell delivery for the animal (saline injection) and may not be possible for humans since it may not be ethical to leave humans untreated. The third row is "Standard of care," which is not applicable for cell culture, but could be an existing chemotherapy for animals and humans. The fourth row is "Unmodified T-cells" only, which can establish if CAR expression is required for the MOA. The fifth row is "CAR T-cells with empty CAR," which can establish if a specific CAR must be expressed or if any CAR can suffice. The sixth row is "Dead CAR

T-cells" which tests if live cells are required. Finally, the seventh row is "Fibroblasts expressing CAR" to determine if T-cells are required or if any cell can suffice. Note that most experiments in Table 3 can be done in cells or animal models, but, due to ethical considerations, only one or two of the experiments can be done in CTP clinical trials. This example application highlights the difficulty in establishing MOA in humans. Clinical trials are designed to assess efficacy and safety, not to establish MOA or validate a potency test.

Example application: establishing MOA for a tissue engineered medical product
Another example application is given in Table 4: establishing the MOA of a tissue engineered medical product composed of autologous cultured chondrocytes on a collagen membrane for treating knee cartilage defects (based on MACI). The proposed MOA is cartilage regeneration by the following activities: i) the implant will occupy the trauma site to prevent scar formation, ii) the cells will proliferate and secrete cartilage matrix that will form new cartilage tissue, iii) the scaffold will help retain the cells at implantation site providing a supportive niche for the cells to proliferate, differentiate and make new cartilage and iv) the cells will secrete factors to recruit other cells to the implantation site to support cartilage regeneration.

The test article is "Cells+scaffold" as shown in the first row of Table 4. The second row is "Negative control," to assess if the test article has an effect above background levels. The negative control could be an untreated well for cell culture, a sham surgery for animals and may not be applicable for humans (since it may not be ethical to leave humans untreated). The third row is "Standard of care," which is not applicable for cell culture, but could be microfracture for animals and humans. The fourth row is "Scaffold" only, which can establish if cells are required for the MOA. The fifth row is "Cells" only, which can establish if the cells by themselves can regenerate cartilage or if the scaffold is required. The sixth and seventh rows use dead cells instead of live cells, to establish that live cells are required for cartilage regeneration. The eighth and ninth rows use non-relevant cells, such as fibroblasts, to establish that the chondrocytes are required for the MOA. The tenth and eleventh rows use scaffolds with the wrong porosity (out of specification), to establish that scaffolds with the specified porosity are required for the MOA. All the experiments could be conducted in cells or animals, but only one, maybe two, can be conducted in humans. These examples demonstrate how clinical trials are not designed to establish the product's MOA.

Table 3 Example application for establishing MOA for a hypothetical product: chimeric antigen receptor (CAR) T-cell therapy directed against leukemia cells expressing a target surface antigen (based on Kymriah)

Treatment	What does it test?	Cell Culture: Assess CAR T-cell killing of cells expressing the target surface antigen	Animal Model: Treat animals with leukemia	Human Clinical Trial: Treat leukemia patients & assess remission (by lymphoblasts in bone marrow, blasts in peripheral blood, extramedullary disease, neutrophils, platelets)
1) CAR T-cells directed against leukemia cells expressing a target surface antigen	This is the test article	✓	✓	✓
2) Negative control (placebo, sham)	Does test article have an effect greater than natural healing response?	✓ (untreated dish of cells expressing the target surface antigen)	✓ (sham infusion)	*May not be ethical*
3) Standard of care (chemotherapy)	How does test article compare with current treatment options?	*not applicable*	*May not be an existing treatment*	*May not be an existing treatment*
4) Unmodified T cells	Is the CAR necessary? (Can T-cells work alone?)	✓	✓	**X**
5) CAR T-cells with empty CAR (CAR lacks binding site for leukemia cell receptor)	Is a specific CAR required or can any CAR suffice? (can CAR work without being directed towards leukemia cell surface antigen?)	✓	✓	**X**
6) Dead CAR T-cells directed against leukemia cells expressing a target surface antigen	Are live cells required? (Can dead cells suffice?)	✓	✓	**X**
7) Fibroblasts expressing CAR directed against leukemia cells expressing a target surface antigen	Are T-cells required? (Will other cell types work?)	✓	✓	**X**

*The first column in each row indicates control experiments that could be performed to establish the product's MOA. Grey shading in the first row indicates the test article, which is the product in its intended form. The second column of each row indicates what each experiment would test. Three different test systems are given in the third column (cell culture), fourth column (animal model) and fifth column (human clinical trial). Green check marks indicate that an experiment can be performed in the given test system while red X indicates that it is not appropriate to conduct the experiment in the test system. All clinical trials for the 27 US FDA approved CTPs conducted to date are either single-arm or two-arm trials. The experiments described in the first column are meant to serve as examples and are not meant to serve as recommendations for experiments that should be conducted to establish the MOA of a CAR T-cell therapy.

Table 4 Example application for establishing MOA for a hypothetical product: Autologous cultured chondrocytes on collagen membrane for treating knee cartilage defects (based on MACI)

Treatment	What does it test?	Cell Culture: Assess cartilage formation in cell culture dish using histology	Animal Model: Implant in animal with cartilage defect & assess regeneration by histology	Human Clinical Trial: Implant in patients with knee cartilage defect & assess cartilage repair using a questionnaire (KOOS)
1) Cells + scaffold	This is the test article	✓	✓	✓
2) Negative Control (placebo, sham)	Does test article have an effect greater than natural healing response?	✓ (dish with medium only)	✓ (sham surgery)	*not applicable*
3) Standard of care (microfracture)	How does test article compare with current treatment options?	*not applicable*	✓	✓
4) Scaffold	Can scaffold work alone? (are cells required?)	✓	✓	✗
5) Cells	Can cells work alone? (is scaffold required?)	✓	✓	✗
6) Dead cells	Are live cells required? (can dead cells suffice?)	✓	✓	✗
7) Dead cells + scaffold		✓	✓	✗
8) Non-relevant cells (fibroblasts)	Are chondrocytes required? (can any cell suffice?)	✓	✓	✗
9) Non-relevant cells (fibroblasts) + scaffold		✓	✓	✗
10) Scaffold w/wrong porosity	Are scaffold specs meaningful? (can out-of-spec scaffolds suffice?)	✓	✓	✗
11) Cells + Scaffold w/wrong porosity		✓	✓	✗

The first column in each row indicates control experiments that could be performed to establish the product's MOA. Grey shading in the first row indicates the test article, which is the product in its intended form. The second column of each row indicates what each experiment would test. Three different test systems are given in the third column (cell culture), fourth column (animal model) and fifth column (human clinical trial). Green check marks indicate that an experiment can be performed in the given test system while red X indicates that it is not appropriate to conduct the experiment in the test system. All clinical trials for the 27 US FDA approved CTPs conducted to date are either single-arm or two-arm trials. The experiments described in the first column are meant to serve as examples and are not meant to serve as recommendations for experiments that should be conducted to establish the MOA of a cartilage tissue therapy. (KOOS = Knee injury and Osteoarthritis Outcome Score)

Simon Jr. *et al. Journal of Translational Medicine*   (2024) 22:416

Page 12 of 17

Clinical trial structures for 27 CTPs approved in US

Table 2 (fourth column) shows the clinical trial structures (i.e., designs) for the 27 CTPs approved in the US. Only two of these 27 trials provide relevant insight on the MOA. Twenty one of the 27 trials (78%) were single-arm which cannot substantively inform the MOA. Six trials (22%) were two-arm, of which four did not shed much light on MOA: one compared to placebo (Laviv) and three were compared to standard of care (Gintuit, MACI, Stratagraft). One-arm trials, as well as two-arm trials that compare to placebo, untreated or standard of care, are useful for assessing efficacy, but shed little light on MOA. Thus, 25 of the 27 CTP trials (93%) did not substantively inform the MOA.

However, two of the 27 trials (7%) shed light on the MOA. The Provenge trial compared the test article, activated peripheral blood mononuclear cells (PBMCs), to unactivated PBMCs. This tested whether activated PBMCs were required for efficacy. Omisirge, umbilical cord blood (UCB) cells expanded ex vivo in the presence of nicotinamide, was compared to standard of care, which was untreated UCB cells. This tested whether cells expanded in the presence of nicotinamide performed comparably to unmanipulated UCB cells. Thus, only two of the clinical trials for the 27 approved CTPs had any real bearing on the MOA in humans.

Control experiments in a research paper

To take the examples further, consider the challenge of achieving the evidence in a CTP clinical trial that would be required for a peer-reviewed basic research publication. Imagine if a research paper were submitted that reported only one experiment, the "cells + scaffold" experiment from Table 4, to assess cartilage regeneration in an animal model. Furthermore, imagine for this hypothetical study that no controls were run: no negative control, no cells only, no scaffold only, and no microfracture. The manuscript would likely be rejected with the reviewers requiring that the control experiments be conducted. Applying this same standard is neither feasible nor ethically responsible for CTP clinical trials, further illustrating the difficulty in establishing MOA in clinical trials.

CTP clinical trials are designed to assess efficacy—not establish MOA

Given that only single-arm or two-arm clinical trials are conducted for CTPs, it may be unrealistic at this time to think that the MOA of a CTP could be established in humans with certainty. Indeed, CTP clinical trials are not designed to establish MOA but are instead designed to assess safety and efficacy. Even for cells and animals, where many types of experiments can be conducted (Tables 3 and 4), establishing the MOA is largely

aspirational, as discussed above in the aspirin example. Aspirin affects at least five biochemical pathways which leads to uncertainty regarding its efficacy-relevant MOA in cells or animals.

Considerations for the development of potency tests

Potency test validation

Careful validation procedures for characterizing potency test performance are critical for assay reliability. Potency tests should be validated for specificity, sensitivity, linearity, range, accuracy, precision, repeatability (within lab variability), reproducibility (between lab variability), detection limit, quantification limit, robustness and fit for purpose [18, 59–61]. It should be established that the potency test can perform within defined specifications to help assure consistency in the manufactured product. Those specifications are set by measuring reference samples under a variety of conditions, modifying the assay to reduce uncertainty, and qualifying the assay response with respect to its precision and accuracy. The resulting criteria that are established for precision and accuracy in measurement of the reference sample are specifications that must be met when running the assay with the reference sample in parallel with the CTP sample.

Matrix approach to potency tests

Use of several different potency tests, or a potency assay matrix [62], may be useful for CTPs for several reasons. First, CTPs likely have multiple MOAs. Second, there may be multiple product attributes that could be used as a potency attribute for any one of the MOAs for a given product. As discussed for a CAR T-cell therapy, its potency attribute could be defined as its ability to secrete IFN-γ upon binding to target cells (Fig. 2b), its ability to secrete IL5 upon binding to target cells, the ability of the CAR T-cells to kill target cells or a complex response that involves all three activities. Third, there may be a variety of ways to measure a given potency attribute. As discussed above, CAR T-cell release of IFN-γ release could be measured by ELISA or Western blot.

An assay matrix is also useful for the application of orthogonal methods for measuring a potency attribute in order to improve measurement confidence. Orthogonal measurements use different physical principles to measure the same property of the same sample with the goal of detecting method-specific biases and interferences [63]. The comparison of orthogonal methods can establish confidence in the accuracy of an assay. An example could be measurement of the ability of CAR T-cells to kill CD19-expressing cells by chromium release of target cells and by impedance plates (Fig. 2c). These measurements are orthogonal to one another because they are

based on different physical principles: chromium release monitors cell health by assessing membrane integrity while impedance plates monitor cell health by the cells' ability to adhere to the plate (since impedance decreases when cells detach). When a similar measurement result is obtained from the two techniques, confidence in the results is enhanced.

Potency test variability, product variability, efficacy endpoint test variability and patient variability

It is difficult to determine if variability in potency test results is due to variability inherent in the test measurement, in the product or both. There may be variability in the biological reagents used for the potency test, such as the target cells in an IFN-γ release assay or in a cell-killing assay. There is well-known variability in antibody-based detection, such as detection of a secreted protein, like IFN-γ, by an ELISA. Assay validation strives to control for these variables with additional measurements that establish criteria for reagent storage and handling. There may be variability in the CTP, both within a batch or between batches.

Another challenge is variability in the patient population receiving the treatment. Patients in a population may vary from one another due to inherent biological, environmental and clinical factors, including previous treatments. Each patient's condition may require different functionalities (*i.e.,* MOAs) from the product. Finally, there is variability in the efficacy endpoint testing, such as determining the amount of lymphoblasts in bone marrow [30] for a CAR T-cell clinical trial, that make it challenging to assess efficacy. Reducing variability in the product itself, the product manufacturing process, product testing (e.g., potency test), and efficacy endpoint tests, makes it more likely that a statistically significant clinical benefit to patients can be detected in a clinical trial. Reducing patient variability, by narrowing the indication, or refining the inclusion and exclusion criteria, may also increase the likelihood of detecting efficacy in a clinical trial.

Short product shelf life

Short product shelf life may limit the availability of potency test results at the point of product release if the product has a shelf life that is shorter than the time it takes to conduct the potency test. These cases require rapid alternative methods of obtaining equivalent data for release. If a compendial method is required and that assay exceeds the shelf-life, a two-stage release may be appropriate using an alternative rapid assay for initial release followed by confirmation of the batch disposition by a validated, longer-term test.

Stability and manufacturing changes

A potency test ought to be sensitive to product stability [7], such that products that have degraded from aging should demonstrate low potency test results. When changes are made to a manufacturing process, potency tests are important for assessing comparability between newer and older lots to assess how the changes may affect product performance [64].

Potency assurance strategy

FDA recently released a new draft guidance for a potency assurance strategy for cell and gene therapy products [8]. "A potency assurance strategy is a multifaceted approach that reduces risks to the potency of a product through manufacturing process design, manufacturing process control, material control, in-process testing and potency lot release assays." The new guidance emphasizes a life-cycle approach to potency that is grounded in quality risk management, where potency tests are considered throughout the product lifecycle from product development to licensure [65]. A lifecycle approach could allow the potency tests to change during product development as knowledge of the MOA, potency tests and risks to product potency is gained.

Standards for potency tests

An ASTM "Standard Guide for Cell Potency Assays for Cell Therapy and Tissue Engineered Products" was published in 2019 that summarized current perspectives on potency from FDA Guidance Documents, US Pharmacopeia (USP), International Conference on Harmonization (ICH) and European Medicines Agency (EMA) [61]. A feasibility study on potency published in 2022 found that potency tests may not yet be ready for standards development [5]. However, some assays may be ready for development into standard test methods [66], such as CAR T-cell potency tests, cell viability tests and methods for quantifying the fraction of viable stems cells in a cell preparation. Of the six CAR T-cell therapies on the market, three of them reveal IFN-γ secretion and CAR expression as potency tests. In addition, 16 of the 27 approved cell therapies (59%) cite cell viability tests as a potency test (Table 2). These numbers may be higher, since many potency tests are redacted. The use of a cell viability test as a CTP potency test may not be ideal, since cell viability tests may not be specific enough with regard to the MOA. There are many CAR T-cell therapies and other types of cell therapies under development that might benefit from standard test methods for IFN-γ secretion, CAR expression and cell viability. Finally, methods for quantifying the specific fraction of viable tissue stem cells in cell preparations are under

development as standard test methods for assessing the potency of stem cell therapy products [67]. Standard test methods should be vetted by interlaboratory studies to identify sources of variability and to assess reproducibility when the tests are conducted in different labs, by different operators and with different equipment [68].

The future of MOA and potency tests

New methods for determining the MOA of CTPs may be developed in the future. The biological activity assessed in potency tests may be the result of complex interactions. Multi-omics is a promising approach where many omics modalities are used to collect data from the product and patients such as genomics, transcriptomics, proteomics, secretomics, metabolomics and lipidomics [69–71]. Design of experiments and multifactorial experimental designs hold promise for determining how complex biological processes operate [72, 73]. Machine learning and artificial intelligence may be useful for determining MOA and potency attributes from omics data. Systems level thinking, such as computational systems modeling, which takes a top-down approach that focuses on the macroscopic behavior of complex systems to predict behavior, may be useful for predicting the non-linear and emergent behavior of biological systems [74]. Automation and robotics for accelerated and higher throughput testing may also be important [75]. Collection of large amounts of product data and patient data during clinical trials will be helpful, so that the relationship between product attributes and patient attributes can be used to improve understanding of MOAs. It would be helpful if some of these efforts occur in the public domain, so that the data are publicly available [6] (Fig. 1). This may require industrial consortia due to the high costs and amount of effort that are required for clinical studies. Public domain product and patient data would allow the data science community at large to participate in developing innovative analytical methods for determining MOAs.

Potency and MOA are useful concepts

Despite the challenges associated with MOA, potency, and efficacy that are discussed herein, potency tests and efforts to establish an MOA are important for CTPs. Human CTP clinical trials should have a rational and scientific underpinning where a sensible MOA is used as the motivation for product development and human investigation. Having an MOA and potency test that can withstand scientific scrutiny assures that human testing is conducted with a sound basis in scientific reasoning. Biologics are used for their biological activity and it makes sense to assess the quality of a CTP using a potency test

that assesses its biological activity. This perspective offers a framework for interpreting what has been experienced for the US-approved CTPs. This perspective does not suggest that clinical trials should have multiple arms that include more of the experiments shown in Tables 3 and 4. The limitations of clinical trials may feel restrictive from a scientific standpoint, but these constraints are necessary to protect human subjects. The field of CTPs is rapidly evolving and the responses to the challenges presented herein will undoubtedly need to evolve as well.

Conclusions

There are several key takeaways from this perspective:

- A measurand is "the quantity or property *intended* to be measured", reminding us that all measurements have false positives and false negatives.
- A product attribute is often a measurand and should be defined independently from the measurement of the attribute.
- Potency ≠ efficacy (ideally, potency test results correlate with efficacy endpoint test results)
- A CTP can be "potent but not efficacious" or "not potent but efficacious." The CTP development goal is to achieve "potent and efficacious."
- Clinical trials are not designed to establish an MOA or validate a potency test; instead, clinical trials are designed to assess efficacy and safety.
- The clinical trials for the 27 US-approved CTPs were one- or two-arm trials; it is challenging to establish the MOA in humans or validate a potency test with any level of certainty with only one or two arms.
- Potency tests and efforts to establish an MOA are essential, since clinical trials should have a rational and scientific basis in a plausible MOA that guides product development and human investigation.

MOA, potency, potency test, efficacy, efficacy endpoint and efficacy endpoint tests should be independently defined to improve clarity during discussions concerning correlations between potency test results and clinical outcomes. Clarification of these independent terms will help to avoid hidden assumptions that result when concepts such as potency and efficacy are conflated. The ideas and observations presented herein may be helpful to product developers for setting realistic goals for understanding a product's MOA and for establishing correlations between potency tests and clinical efficacy.

Abbreviations

ALL	Acute lymphoblastic leukemia
CTP	Cell therapy product
CAR	Chimeric antigen receptor
COX	Cyclooxygenase

Simon Jr. *et al. Journal of Translational Medicine*   (2024) 22:416

Page 15 of 17

ELISA	Enzyme linked immunosorbent assay
EMA	European Medicines Agency
FDA	Food and Drug Administration
IFN-γ	Interferon-γ
ICH	International Conference on Harmonization
MOA	Mechanism of action
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
PAP-GM-CSF	Human prostatic acid phosphatase (PAP) linked to human granulocyte–macrophage colony-stimulating factor
PBMC	Peripheral blood mononuclear cells
PCR	Polymerase chain reaction
SBRA	Summary Basis for Regulatory Action
US	United States
USP	US Pharmacopeia
VIM	Vocabulary of Metrology

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-024-05179-7.

> **Additional file 1.** Key aspects for each of the 27 US-approved cell therapy products are summarized including product name, year approved, sponsor, product description, indication, clinical trial structure, efficacy endpoints, mechanism of action, potency test, comments and references.
>
> **Additional file 2.** Key definitions that are referred to in the main text.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Author details

¹Biosystems and Biomaterials Division, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA. ²United Therapeutics Corporation, Regenerative Medicine Operations, Research Triangle Park, NC, USA. ³Association for the Advancement of Blood and Biotherapies (AABB), Bethesda, MD, USA. ⁴Niba Labs, Ljubljana, Slovenia. ⁵National Institute of Biology, Ljubljana, Slovenia. ⁶Department of Pediatrics, Children's Healthcare of Atlanta, Marcus Center for Cellular and Gene Therapies, Correlative Studies Laboratory, Emory University School of Medicine, Atlanta, GA, USA. ⁷Pharma Services, Science and Technology, Thermo Fisher Scientific, San Diego, CA, USA. ⁸Medical Education, Sports Medicine and Orthobiologics, Medical Sales Institute, San Diego, CA, USA. ⁹Cellular Therapy Lab, Research Institute of McGill University Health Center, Montreal, QC, Canada. ¹⁰Synthasome, Inc, San Diego, CA, USA. ¹¹Asymmetrex, LLC, Boston, MA, USA. ¹²International Stem Cell Banking Initiative, Barley, Herts, UK. ¹³National Stem Cell Resource Centre, Institute of Zoology, Chinese Academy of Sciences, Beijing, China. ¹⁴Beijing Institute for Stem Cells and Regenerative Medicine, Chinese Academy of Sciences, Beijing, China. ¹⁵Kendall Innovations, Cambridge, MA, USA. ¹⁶Cell and Gene Therapy Facility, Health Sciences Authority, Singapore, Singapore. ¹⁷Science, Research and Innovation Group, Biotherapeutics and Advanced Therapies Division, Medicines and Healthcare Products Regulatory Agency, South Mimms, Hertfordshire, UK.

References

1. Bravery CA, Carmen J, Fong T, Oprea W, Hoogendoorn KH, Woda J, et al. Potency assay development for cellular therapy products: an ISCT review of the requirements and experiences in the industry. Cytotherapy. 2013;15:9-19.e9.
2. de Wolf C, van de Bovenkamp M, Hoefnagel M. Regulatory perspective on in vitro potency assays for human mesenchymal stromal cells used in immunotherapy. Cytotherapy. 2017;19:784–97.
3. Sebastião MJ, Serra M, Gomes-Alves P, Alves PM. Stem cells characterization: OMICS reinforcing analytics. Curr Opin Biotechnol. 2021;71:175–81.
4. Welch AR. CGT Potency Assays: When Will We Wake Up From The Nightmare? Cell & Gene Collaborative, 2022. Accessed February 15, 2024: https://www.cellandgene.com/doc/cell-and-gene-therapy-potency-assays-when-will-we-wake-up-from-the-nightmare-0001?utm_source=Onesignal&utm_medium=Pushnotification&utm_term=Assays.
5. Standards Coordinating Body. Assessing Feasibility of a Standard for Product Potency and Functionality Measurement Methods, 2022. Accessed February 15, 2024: https://www.standardscoordinatingbody.org/feasibility-assessments.
6. Alliance for Regenerative Medicine. Addressing potency-assay related development delays for cell and gene therapies: Results of a scientific exchange between FDA and developers, 2023. Accessed February 14, 2024: https://alliancerm.org/potency-assay-report/.
7. Guidance for Industry—Potency Tests for Cellular and Gene Therapy Products, US Food and Drug Administration, 2011. Accessed February 15, 2024: https://www.fda.gov/media/79856/download.
8. Draft Guidance for Industry—Potency Assurance for Cellular and Gene Therapy Products, US Food and Drug Administration, 2023. Accessed February 14, 2024: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/potency-assurance-cellular-and-gene-therapy-products.
9. Approved Cellular and Gene Therapy Products, US Food and Drug Administration. Accessed February 15, 2024: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/approved-cellular-and-gene-therapy-products.
10. Slides for the July 12, 2017 Meeting of the Oncologic Drugs Advisory Committee (ODAC). Novartis Slides, FDA, 2017. Accessed February 15, 2024: https://www.fda.gov/media/106496/download.
11. FDA Briefing Document—Oncologic Drugs Advisory Committee Meeting, Kymriah, 2017. Accessed February 15, 2024: https://www.fda.gov/media/106081/download.

Simon Jr. *et al. Journal of Translational Medicine*   (2024) 22:416

Page 16 of 17

12. International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM), 3rd ed.; Joint Committee for Guides in Metrology (JCGM), 2012. Accessed February 15, 2024: https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf.

13. Arora D, Babakhanova G, Simon CG. Tissue engineering measurands. ACS Biomater Sci Eng. 2020;6:5368–76.

14. Plant AL, Locascio LE, May WE, Gallagher PD. Improved reproducibility by assuring confidence in measurements in biomedical research. Nat Methods. 2014;11:895–8.

15. Guidance for Industry and Review Staff—Labeling for Human Prescription Drug and Biological Products—Determining Established Pharmacologic Class for Use in the Highlights of Prescribing Information - Good Review Practice. US Food and Drug Administration, 2009. Accessed February 15, 2024: https://www.fda.gov/media/77834/download.

16. 21 CFR 3.2(k), Code of Federal Regulations. US National Archives. Accessed February 15, 2024: https://www.ecfr.gov/current/title-21/chapter-I/subchapter-A/part-3/subpart-A/section-3.2.

17. 21 CFR 600.3(s), Code of Federal Regulations, US National Archives. Accessed February 15, 2024: https://www.ecfr.gov/current/title-21/chapter-I/subchapter-F/part-600/subpart-A/section-600.3#p-600.3(s).

18. International Conference on Harmonisation (ICH). ICH Harmonised Tripartite Guideline Specifications: Test Procedures And Acceptance Criteria For Biotechnological/Biological Products Q6B (ICHQ6B), Current Step 4 version, 1999. Accessed February 15, 2024: https://www.ich.org/page/quality-guidelines.

19. ISO 23033—Biotechnology—Analytical methods—General requirements and considerations for the testing and characterization of cellular therapeutic products. International Organization for Standardization: Geneva, Switzerland, 2021.

20. 21 CFR 610.10, Code of Federal Regulations, US National Archives. Accessed February 15, 2024: https://www.ecfr.gov/current/title-21/chapter-I/subchapter-F/part-610/subpart-B/section-610.10.

21. Guidance for Industry—Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. Food and Drug Administration, 1998. Accessed February 15, 2024: https://www.fda.gov/files/drugs/published/Providing-Clinical-Evidence-of-Effectiveness-for-Human-Drug-and-Biological-Products.pdf.

22. Draft Guidance for Industry—Multiple Endpoints in Clinical Trials, US Food and Drug Administration, 2017. Accessed February 15, 2024: https://www.fda.gov/files/drugs/published/Multiple-Endpoints-in-Clinical-Trials-Guidance-for-Industry.pdf.

23. Surrogate Endpoint Resources for Drug and Biologic Development, US Food and Drug Administration. Accessed February 15, 2024: https://www.fda.gov/drugs/development-resources/surrogate-endpoint-resources-drug-and-biologic-development.

24. Schreiber RD. Measurement of mouse and human interferon γ. Curr Protocols Immunol. 2000;37:681–8.

25. Harrison P, Briggs C, Machin SJ. Platelet counting. Methods Mol Biol. 2004;272:29–46.

26. Brunck MEG, Andersen SB, Timmins NE, Osborne GW, Nielsen LK. Absolute counting of neutrophils in whole blood using flow cytometry. Cytometry A. 2014;85:1057–64.

27. Summary Basis for Regulatory Action—Kymriah, FDA, 2017. Accessed February 15, 2024: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/kymriah-tisagenlecleucel.

28. Package Insert—Kymriah, FDA, 2017. Accessed February 15, 2024: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/kymriah-tisagenlecleucel.

29. Maude SL, Laetsch TW, Buechner J, Rives S, Boyer M, Bittencourt H, et al. Tisagenlecleucel in children and young adults with B-cell lymphoblastic leukemia. N Engl J Med. 2018;378:439–48.

30. Gupta S, Devidas M, Loh ML, Raetz EA, Chen S, Wang C, et al. Flow-cytometric vs. -morphologic assessment of remission in childhood acute lymphoblastic leukemia: a report from the Children's Oncology Group (COG). Leukemia. 2018;32:1370–9.

31. Xhangolli I, Dura B, Lee G, Kim D, Xiao Y, Fan R. Single-cell analysis of CAR-T cell activation reveals a mixed TH1/TH2 response independent of differentiation. Genom Proteom Bioinform. 2019;17:129–39.

32. Kiesgen S, Messinger JC, Chintala NK, Tano Z, Adusumilli PS. Comparative analysis of assays to measure CAR T-cell-mediated cytotoxicity. Nat Protoc. 2021;16:1331–42.

33. Rapko S, Parker A, Mortelliti C, Duguay SJ. #P192—Aggrecan gene expression as a potency marker for matrix-induced autologous chondroctye implantation (MACI). Osteoarthritis and Cartilage 15, Supplement B, page B136, 2007. https://doi.org/10.1016/S1063-4584(07)61547.

34. Saris D, Price A, Widuchowski W, Bertrand-Marchand M, Caron J, Drogset JO, et al. Matrix-applied characterized autologous cultured chondrocytes versus microfracture: two-year follow-up of a prospective randomized trial. Am J Sports Med. 2014;42:1384–94.

35. Summary Basis for Regulatory Action—MACI, FDA, 2016. Accessed February 15, 2024: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/maci-autologous-cultured-chondrocytes-porcine-collagen-membrane.

36. Package Insert—MACI, FDA, 2016. Accessed February 15, 2024: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/maci-autologous-cultured-chondrocytes-porcine-collagen-membrane.

37. Fändriks A, Tranberg R, Karlsson J, Möller M, Zügner R. Gait biomechanics in patients with intra-articular tibial plateau fractures—gait analysis at three months compared with age- and gender-matched healthy subjects. BMC Musculoskelet Disord. 2021;22:702.

38. List of Cleared or Approved Companion Diagnostic Devices (In Vitro and Imaging Tools), US Food and Drug Administration. Accessed February 15, 2024: https://www.fda.gov/medical-devices/in-vitro-diagnostics/list-cleared-or-approved-companion-diagnostic-devices-in-vitro-and-imaging-tools.

39. Krejsa C, Rogge M, Sadee W. Protein therapeutics: new applications for pharmacogenetics. Nat Rev Drug Discov. 2006;5:507–21.

40. Package Insert—Aspirin, FDA, 2015. Accessed February 15, 2024: https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/200671s000lbl.pdf.

41. Janeway CA Jr, Travers P, Walport M & Shlomchik MJ. T cell-mediated cytotoxicity. In Immunobiology: the immune system in health and disease, 5th edition. New York: Garland Science; 2001. T cell-mediated cytotoxicity. Accessed October 17, 2023: https://www.ncbi.nlm.nih.gov/books/NBK27101/.

42. Vane JR, Botting RM. The mechanism of action of aspirin. Thromb Res. 2003;110:255–8.

43. Vane JR. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. Nat New Biol. 1971;231:232–5.

44. Hemler M, Lands WE. Purification of the cyclooxygenase that forms prostaglandins. Demonstration of two forms of iron in the holoenzyme. J Biol Chem. 1976;251:5575–9.

45. Roth GJ, Majerus PW. The mechanism of the effect of aspirin on human platelets. I. Acetylation of a particulate fraction protein. J Clin Invest. 1975;56:624–32.

46. The Nobel Prize, John R. Vane. Accessed February 15, 2024: https://www.nobelprize.org/prizes/medicine/1982/vane/facts/.

47. Xie WL, Chipman JG, Robertson DL, Erikson RL, Simmons DL. Expression of a mitogen-responsive gene encoding prostaglandin synthase is regulated by mRNA splicing. Proc Natl Acad Sci. 1991;88:2692–6.

48. Chandrasekharan NV, Dai H, Roos KLT, Evanson NK, Tomsik J, Elton TS, et al. COX-3, a cyclooxygenase-1 variant inhibited by acetaminophen and other analgesic/antipyretic drugs: cloning, structure, and expression. Proc Natl Acad Sci. 2002;99:13926–31.

49. Somasundaram, Sigthorsson, Simpson, Watts, Jacob, Tavares, et al. Uncoupling of intestinal mitochondrial oxidative phosphorylation and inhibition of cyclooxygenase are required for the development of NSAID-enteropathy in the rat. Aliment Pharmacol Ther. 2000;14:639–50.

50. Paul-Clark MJ, van Cao T, Moradi-Bidhendi N, Cooper D, Gilroy DW. 15-epi-lipoxin A4–mediated induction of nitric oxide explains how aspirin inhibits acute inflammation. J Exp Med. 2004;200:69–78.

51. McCarty MF, Block KI. Preadministration of high-dose salicylates, suppressors of NF-κB activation, may increase the chemosensitivity of many cancers: an example of proapoptotic signal modulation therapy. Integr Cancer Ther. 2006;5:252–68.

52. Bertolini A, Ferrari A, Ottani A, Guerzoni S, Tacchi R, Leone S. Paracetamol: new vistas of an old drug. CNS Drug Rev. 2006;12:250–75.

53. Package Insert—Acetaminophen, US Food and Drug Administration, 2015. Accessed February 16, 2024: https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/204767s000lbl.pdf.

Simon Jr. *et al. Journal of Translational Medicine*   (2024) 22:416

Page 17 of 17

54. Flower RJ, Vane JR. Inhibition of prostaglandin synthetase in brain explains the anti-pyretic activity of paracetamol (4-acetamidophenol). Nature. 1972;240:410–1.
55. Zygmunt PM, Chuang H, Movahed P, Julius D, Högestätt ED. The anandamide transport inhibitor AM404 activates vanilloid receptors. Eur J Pharmacol. 2000;396:39–42.
56. Ottani A, Leone S, Sandrini M, Ferrari A, Bertolini A. The analgesic activity of paracetamol is prevented by the blockade of cannabinoid CB1 receptors. Eur J Pharmacol. 2006;531:280–1.
57. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov. 2006;5:993–6.
58. Davis RL. Mechanism of action and target identification: a matter of timing in drug discovery. iScience. 2020;23:101487.
59. Ederveen J. A practical approach to biological assay validation. Dutch Ministry of Housing, Spatial Planning and the Environment (VROM), 2010. Accessed February 15, 2024: https://edraservices.nl/wp-content/uploads/2022/02/Assay-validation-P08090.pdf.
60. FDA Guidance—Bioanalytical Method Validation—Guidance for Industry, U.S. Food and Drug Administration, 2018. Accessed July 31, 2022: https://www.fda.gov/media/70858/download.
61. ASTM F3368 Standard Guide for Cell Potency Assays for Cell Therapy and Tissue Engineered Products; ASTM International: West Conshohocken, PA, 2019.
62. Chinnadurai R, Rajan D, Qayed M, Arafat D, Garcia M, Liu Y, et al. Potency analysis of mesenchymal stromal cells using a combinatorial assay matrix approach. Cell Rep. 2018;22:2504–17.
63. Simon CG, Borgos SE, Calzolai L, Nelson BC, Parot J, Petersen EJ, et al. Orthogonal and complementary measurements of properties of drug products containing nanomaterials. J Control Release. 2023;354:120–7.
64. Draft Guidance for Industry Manufacturing Changes and Comparability for Human Cellular and Gene Therapy Products, US Food and Drug Administration, 2023. Accessed February 15, 2024: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/manufacturing-changes-and-comparability-human-cellular-and-gene-therapy-products.
65. Guidance for Industry: Q9 (R1) Quality Risk Management. US Food and Drug Administration, 2023. Accessed February 16, 2024: https://www.fda.gov/media/167721/download.
66. Wilson AJ, Brown N, Rand E, Genever PG. Attitudes towards standardization of mesenchymal stromal cells-a qualitative exploration of expert views. Stem Cells Transl Med. 2023;12:745.
67. Chopra H, Daley MP, Kumar A, Sugai J, Dahlkemper A, Kaigler D, et al. Evaluation of the precision of kinetic stem cell (KSC) counting for specific quantification of human mesenchymal stem cells in heterogeneous tissue cell preparations. Life. 2024;14:51.
68. Simon CG, Lin-Gibson S, Elliott JT, Sarkar S, Plant AL. Strategies for achieving measurement assurance for cell therapy products. Stem Cells Transl Med. 2016;5:705–8.
69. Bayne EF, Simmons AD, Roberts DS, Zhu Y, Aballo TJ, Wancewicz B, et al. Multiomics method enabled by sequential metabolomics and proteomics for human pluripotent stem-cell-derived cardiomyocytes. J Proteome Res. 2021;20:4646–54.
70. Odeh-Couvertier VY, Dwarshuis NJ, Colonna MB, Levine BL, Edison AS, Kotanchek T, et al. Predicting T-cell quality during manufacturing through an artificial intelligence-based integrative multiomics analytical platform. Bioeng Transl Med. 2022;7: e10282.
71. Sakallioglu IT, Maroli AS, De Lima LA, Marshall DD, Evans BW, Zinniel DK, et al. Multi-omics investigation into the mechanism of action of an anti-tubercular fatty acid analogue. J Am Chem Soc. 2022;144:21157–73.
72. Bukys MA, Mihas A, Finney K, Sears K, Trivedi D, Wang Y, et al. High-dimensional design-of-experiments extracts small-molecule-only induction conditions for dorsal pancreatic endoderm from pluripotency. iScience. 2020;23:101346.
73. Sears KE, Gullapalli K, Trivedi D, Mihas A, Bukys MA, Jensen J. Controlling neural territory patterning from pluripotency using a systems developmental biology approach. iScience. 2022;25:104133.
74. National Academies of Sciences, Engineering, and Medicine. 2021. Applying Systems Thinking to Regenerative Medicine: Proceedings of a Workshop. Washington, DC: The National Academies Press. https://doi.org/10.17226/26025.
75. Hillson N, Caddick M, Cai Y, Carrasco JA, Chang MW, Curach NC, et al. Building a global alliance of biofoundries. Nat Commun. 2019;10:2040.

Publisher's Note